

A-Geneseq

! FINDPATTERNS on geneseqp:* allowing 0 mismatches

! 1
E(R,Q,S,H,Y,E) (G,S,Q,I,T,N,P) (D,F,T,S,P,L,I) (N,S,K,N,T,D) (N,D,K,T,E) (L,F,R,I) (T,
S,N)R(I,V,A) October 5, 2004 12:51 ..

AAR22511 ck: 2459 len: 488 ! Aar22511 Human Factor Xai. 7/1992

1
E(R,Q,S,H,Y,E) (G,S,Q,I,T,N,P) (D,F,T,S,P,L,I) (N,S,K,N,T,D) (N,D,K,T,E) (L,F,R,I) (T,
S,N)R(I,V,A)

E(R) (G) (D) (N) (N) (L) (T) R(I)
226: NQTQP ERGDNNLTRI VGGQE

AAR22512 ck: 2469 len: 488 ! Aar22512 Mutated precursor of human
Factor X analogue. 7/1992

1
E(R,Q,S,H,Y,E) (G,S,Q,I,T,N,P) (D,F,T,S,P,L,I) (N,S,K,N,T,D) (N,D,K,T,E) (L,F,R,I) (T,
S,N)R(I,V,A)

E(R) (G) (D) (N) (N) (L) (T) R(I)
226: NQTQP ERGDNNLTRI VGGQE

AAR30729 ck: 500 len: 870 ! Aar30729 p100 protein from human
herpes virus type 6. 8/2003

1
E(R,Q,S,H,Y,E) (G,S,Q,I,T,N,P) (D,F,T,S,P,L,I) (N,S,K,N,T,D) (N,D,K,T,E) (L,F,R,I) (T,
S,N)R(I,V,A)

E(S) (Q) (F) (S) (D) (I) (N) R(A)
120: NKEKF ESQFSDINRA LLRLG

AAR42456 ck: 3609 len: 543 ! Aar42456 Enzyme involved in
eicosapentaenoic acid (EPA) synthesis. 3/2003

1
E(R,Q,S,H,Y,E) (G,S,Q,I,T,N,P) (D,F,T,S,P,L,I) (N,S,K,N,T,D) (N,D,K,T,E) (L,F,R,I) (T,
S,N)R(I,V,A)

E(Q) (Q) (L) (K) (D) (F) (S) R(A)
30: DVQVM EQQLKDFSRA CYVVN

AAR35762 ck: 40 len: 448 ! Aar35762 Factor X (X). 3/2003

1
E(R,Q,S,H,Y,E) (G,S,Q,I,T,N,P) (D,F,T,S,P,L,I) (N,S,K,N,T,D) (N,D,K,T,E) (L,F,R,I) (T,
S,N)R(I,V,A)

E(R) (G) (D) (N) (N) (L) (T) R(I)
186: NQTQP ERGDNNLTRI VGGQE

AAR99465 ck: 3609 len: 543 ! Aar99465 Biosynthetic enzyme of
icosapentaenoic acid synthase. 1/1997

1
E(R,Q,S,H,Y,E) (G,S,Q,I,T,N,P) (D,F,T,S,P,L,I) (N,S,K,N,T,D) (N,D,K,T,E) (L,F,R,I) (T,
S,N)R(I,V,A)

E(Q)(Q)(L)(K)(D)(F)(S)R(A)
30: DVQVM EQQLKDFSRA CYVVN

AAW05820 ck: 8643 len: 306 ! Aaw05820 Factor X heavy chain.
12/1996

1
E(R,Q,S,H,Y,E) (G,S,Q,I,T,N,P) (D,F,T,S,P,L,I) (N,S,K,N,T,D) (N,D,K,T,E) (L,F,R,I) (T,
S,N)R(I,V,A)

E(R)(G)(D)(N)(N)(L)(T)R(I)
44: NQTQP ERGDNNLTRI VGGQE

AAW76216 ck: 2975 len: 488 ! Aaw76216 Human Factor X protein.
11/1998

1
E(R,Q,S,H,Y,E) (G,S,Q,I,T,N,P) (D,F,T,S,P,L,I) (N,S,K,N,T,D) (N,D,K,T,E) (L,F,R,I) (T,
S,N)R(I,V,A)

E(R)(G)(D)(N)(N)(L)(T)R(I)
226: NQTQP ERGDNNLTRI VGGQE

AAW40283 ck: 1666 len: 467 ! Aaw40283 Human Factor X protease.
6/1998

1
E(R,Q,S,H,Y,E) (G,S,Q,I,T,N,P) (D,F,T,S,P,L,I) (N,S,K,N,T,D) (N,D,K,T,E) (L,F,R,I) (T,
S,N)R(I,V,A)

E(R)(G)(D)(N)(N)(L)(T)R(I)
208: NQTQP ERGDNNLTRI VGGQE

AAW37053 ck: 3609 len: 543 ! Aaw37053 S. putrefaciens EPO
biosynthesis gene cluster ORF9 product. 7/199

1
E(R,Q,S,H,Y,E) (G,S,Q,I,T,N,P) (D,F,T,S,P,L,I) (N,S,K,N,T,D) (N,D,K,T,E) (L,F,R,I) (T,
S,N)R(I,V,A)

E(Q)(Q)(L)(K)(D)(F)(S)R(A)
30: DVQVM EQQLKDFSRA CYVVN

AAW76218 ck: 2975 len: 488 ! Aaw76218 Human Factor X protein.
11/1998

1
E(R,Q,S,H,Y,E) (G,S,Q,I,T,N,P) (D,F,T,S,P,L,I) (N,S,K,N,T,D) (N,D,K,T,E) (L,F,R,I) (T,
S,N)R(I,V,A)

E(R)(G)(D)(N)(N)(L)(T)R(I)
226: NQTQP ERGDNNLTRI VGGQE

AAW66092 ck: 20 len: 448 ! Aaw66092 Human factor X variant.
11/1998

1
E(R,Q,S,H,Y,E) (G,S,Q,I,T,N,P) (D,F,T,S,P,L,I) (N,S,K,N,T,D) (N,D,K,T,E) (L,F,R,I) (T,
S,N)R(I,V,A)
E(R) (G) (D) (N) (N) (L) (T) R(I)
186: NQTQP ERGDNNLTTRI VGGQE

AAW89403 ck: 3609 len: 543 ! Aaw89403 S. putrefaciens PKS-like
cluster ORF 9 protein. 10/1999

1
E(R,Q,S,H,Y,E) (G,S,Q,I,T,N,P) (D,F,T,S,P,L,I) (N,S,K,N,T,D) (N,D,K,T,E) (L,F,R,I) (T,
S,N)R(I,V,A)
E(Q) (Q) (L) (K) (D) (F) (S) R(A)
30: DVQVM EQQLKDFSRA CYVVN

AAB10470 ck: 3609 len: 543 ! Aab10470 Shewanella putrefaciens
PKS protein ORF9. 12/2000

1
E(R,Q,S,H,Y,E) (G,S,Q,I,T,N,P) (D,F,T,S,P,L,I) (N,S,K,N,T,D) (N,D,K,T,E) (L,F,R,I) (T,
S,N)R(I,V,A)
E(Q) (Q) (L) (K) (D) (F) (S) R(A)
30: DVQVM EQQLKDFSRA CYVVN

AAG52216 ck: 7182 len: 274 ! Aag52216 Arabidopsis thaliana
protein fragment SEQ ID NO: 66352. 10/2000

1
E(R,Q,S,H,Y,E) (G,S,Q,I,T,N,P) (D,F,T,S,P,L,I) (N,S,K,N,T,D) (N,D,K,T,E) (L,F,R,I) (T,
S,N)R(I,V,A)
E(E) (N) (S) (N) (T) (L) (N) R(V)
168: QQVNP EENSNTLN RV NLGEQ

AAG52217 ck: 376 len: 217 ! Aag52217 Arabidopsis thaliana
protein fragment SEQ ID NO: 66353. 10/2000

1
E(R,Q,S,H,Y,E) (G,S,Q,I,T,N,P) (D,F,T,S,P,L,I) (N,S,K,N,T,D) (N,D,K,T,E) (L,F,R,I) (T,
S,N)R(I,V,A)
E(E) (N) (S) (N) (T) (L) (N) R(V)
111: QQVNP EENSNTLN RV NLGEQ

AAG52218 ck: 9159 len: 151 ! Aag52218 Arabidopsis thaliana
protein fragment SEQ ID NO: 66354. 10/2000

1
E(R,Q,S,H,Y,E) (G,S,Q,I,T,N,P) (D,F,T,S,P,L,I) (N,S,K,N,T,D) (N,D,K,T,E) (L,F,R,I) (T,
S,N)R(I,V,A)

E(E) (N) (S) (N) (T) (L) (N) R(V)
45: QQVNP EENSNTLNRV NLGEQ

AAG64458 ck: 9499 len: 542 ! Aag64458 *S. putrefaciens*
eicosapentanoic acid synthesis enzyme 6. 10/2001

1
E(R,Q,S,H,Y,E) (G,S,Q,I,T,N,P) (D,F,T,S,P,L,I) (N,S,K,N,T,D) (N,D,K,T,E) (L,F,R,I) (T,
S,N) R(I,V,A)
E(Q) (Q) (L) (K) (D) (F) (S) R(A)
30: DVQVM EQQLKDFSRA CYVVN

AAE12822 ck: 9069 len: 635 ! Aae12822 *Caenorhabditis elegans* 5-
HT3 receptor protein, F18. 1/2002

1
E(R,Q,S,H,Y,E) (G,S,Q,I,T,N,P) (D,F,T,S,P,L,I) (N,S,K,N,T,D) (N,D,K,T,E) (L,F,R,I) (T,
S,N) R(I,V,A)
E(E) (I) (L) (D) (E) (R) (S) R(I)
576: QHEIR EEILDERSRI QWQWQ

AAU39422 ck: 2758 len: 491 ! Aau39422 *Propionibacterium acnes*
immunogenic protein #318. 2/2002

1
E(R,Q,S,H,Y,E) (G,S,Q,I,T,N,P) (D,F,T,S,P,L,I) (N,S,K,N,T,D) (N,D,K,T,E) (L,F,R,I) (T,
S,N) R(I,V,A)
E(Q) (Q) (P) (S) (N) (L) (S) R(V)
12: VAQRI EQQPSNLSRV GSIPT

AAB70411 ck: 2975 len: 488 ! Aab70411 Human factor X protein
sequence SEQ ID NO:2. 5/2001

1
E(R,Q,S,H,Y,E) (G,S,Q,I,T,N,P) (D,F,T,S,P,L,I) (N,S,K,N,T,D) (N,D,K,T,E) (L,F,R,I) (T,
S,N) R(I,V,A)
E(R) (G) (D) (N) (N) (L) (T) R(I)
226: NQTQP ERGDNNLTRI VGGQE

ABM68937 ck: 7539 len: 591 ! Abm68937 *Photobacterium luminescens*
protein sequence #2034. 11/2003

1
E(R,Q,S,H,Y,E) (G,S,Q,I,T,N,P) (D,F,T,S,P,L,I) (N,S,K,N,T,D) (N,D,K,T,E) (L,F,R,I) (T,
S,N) R(I,V,A)
E(E) (G) (L) (D) (D) (F) (T) R(I)
439: IEQLD EEGLDDFTRI RELLG

ABM35941 ck: 2758 len: 491 ! Abm35941 *Propionibacterium acnes*
predicted ORF-encoded polypeptide #617. 1

1
E(R,Q,S,H,Y,E) (G,S,Q,I,T,N,P) (D,F,T,S,P,L,I) (N,S,K,N,T,D) (N,D,K,T,E) (L,F,R,I) (T,
S,N)R(I,V,A)

E(Q) (Q) (P) (S) (N) (L) (S) R(V)
12: VAQRI EQQPSNLSRV GSIPT

ABP96316 ck: 9069 len: 635 ! Abp96316 Caenorhabditis elegans 5-
HT3 receptor protein SEQ ID NO:1. 5/2003

1
E(R,Q,S,H,Y,E) (G,S,Q,I,T,N,P) (D,F,T,S,P,L,I) (N,S,K,N,T,D) (N,D,K,T,E) (L,F,R,I) (T,
S,N)R(I,V,A)

E(E) (I) (L) (D) (E) (R) (S) R(I)
576: QHEIR EEILDERSRI QWQWQ

ADE78992 ck: 53 len: 442 ! Ade78992 Human protein modification
and maintenance molecule (PMMM)-30. 1/

1
E(R,Q,S,H,Y,E) (G,S,Q,I,T,N,P) (D,F,T,S,P,L,I) (N,S,K,N,T,D) (N,D,K,T,E) (L,F,R,I) (T,
S,N)R(I,V,A)

E(R) (G) (D) (N) (N) (L) (T) R(I)
226: NQTQP ERGDNNLTRI VGGQE

ADE78990 ck: 5455 len: 444 ! Ade78990 Human protein modification
and maintenance molecule (PMMM)-28. 1/

1
E(R,Q,S,H,Y,E) (G,S,Q,I,T,N,P) (D,F,T,S,P,L,I) (N,S,K,N,T,D) (N,D,K,T,E) (L,F,R,I) (T,
S,N)R(I,V,A)

E(R) (G) (D) (N) (N) (L) (T) R(I)
182: NQTQP ERGDNNLTRI VGGQE

ADE78991 ck: 9890 len: 377 ! Ade78991 Human protein modification
and maintenance molecule (PMMM)-29. 1/

1
E(R,Q,S,H,Y,E) (G,S,Q,I,T,N,P) (D,F,T,S,P,L,I) (N,S,K,N,T,D) (N,D,K,T,E) (L,F,R,I) (T,
S,N)R(I,V,A)

E(R) (G) (D) (N) (N) (L) (T) R(I)
226: NQTQP ERGDNNLTRI VGGQE

Databases searched:

EMBL, Release 2.0, Released on 29Jan2004, Formatted on 12Feb2004

Total finds:	27
Total length:	282,547,505
Total sequences:	1,586,107
CPU time:	11:25.60

PIR

! FINDPATTERNS on pir:* allowing 0 mismatches

! 1
E(R,Q,S,H,Y,E) (G,S,Q,I,T,N,P) (D,F,T,S,P,L,I) (N,S,K,N,T,D) (N,D,K,T,E) (L,F,R,I) (T,
S,N)R(I,V,A) October 5, 2004 12:51 ..

EXHU ck: 2459 len: 488 ! coagulation factor Xa (EC 3.4.21.6)
precursor [validated] - human

1
E(R,Q,S,H,Y,E) (G,S,Q,I,T,N,P) (D,F,T,S,P,L,I) (N,S,K,N,T,D) (N,D,K,T,E) (L,F,R,I) (T,
S,N)R(I,V,A)

E(R) (G) (D) (N) (N) (L) (T) R(I)
226: NQTQP ERGDNNLTRI VGGQE

WZBE17 ck: 3936 len: 455 ! gene 17 protein - human herpesvirus
3

1
E(R,Q,S,H,Y,E) (G,S,Q,I,T,N,P) (D,F,T,S,P,L,I) (N,S,K,N,T,D) (N,D,K,T,E) (L,F,R,I) (T,
S,N)R(I,V,A)

E(R) (G) (T) (S) (D) (L) (T) R(V)
107: AMNHD ERGTSDLTRV LLSSN

XPBE12 ck: 1901 len: 871 ! major antigenic structural protein
p100 - human herpesvirus 6 (strain U110)

1
E(R,Q,S,H,Y,E) (G,S,Q,I,T,N,P) (D,F,T,S,P,L,I) (N,S,K,N,T,D) (N,D,K,T,E) (L,F,R,I) (T,
S,N)R(I,V,A)

E(S) (Q) (F) (S) (D) (I) (N) R(A)
121: NKEKF ESQFSDINRA LLRLG

S08438 ck: 2803 len: 111 ! vpx protein - human
immunodeficiency virus type 2 D205

1
E(R,Q,S,H,Y,E) (G,S,Q,I,T,N,P) (D,F,T,S,P,L,I) (N,S,K,N,T,D) (N,D,K,T,E) (L,F,R,I) (T,
S,N)R(I,V,A)

E(R) (T) (I) (T) (E) (L) (N) R(V)
25: AFAWL ERTITELNRV AVNHL

H64829 ck: 2176 len: 589 ! ycaO protein - Escherichia coli
(strain K-12)

1
E(R,Q,S,H,Y,E) (G,S,Q,I,T,N,P) (D,F,T,S,P,L,I) (N,S,K,N,T,D) (N,D,K,T,E) (L,F,R,I) (T,
S,N)R(I,V,A)

E(E) (G) (F) (D) (D) (F) (T) R(V)
439: IEQLD EEGFDDFTRV RELLG

D90752 ck: 2189 len: 589 ! hypothetical protein ECs0988
[imported] - Escherichia coli (strain O157:H7

1
E(R,Q,S,H,Y,E) (G,S,Q,I,T,N,P) (D,F,T,S,P,L,I) (N,S,K,N,T,D) (N,D,K,T,E) (L,F,R,I) (T,
S,N)R(I,V,A)

E(E) (G) (F) (D) (D) (F) (T)R(V)
439: IEQLD EEGFDDFTRV RELLG

B85616 ck: 2189 len: 589 ! hypothetical protein ycaO
[imported] - Escherichia coli (strain O157:H7, s

1
E(R,Q,S,H,Y,E) (G,S,Q,I,T,N,P) (D,F,T,S,P,L,I) (N,S,K,N,T,D) (N,D,K,T,E) (L,F,R,I) (T,
S,N)R(I,V,A)

E(E) (G) (F) (D) (D) (F) (T)R(V)
439: IEQLD EEGFDDFTRV RELLG

AE0613 ck: 3304 len: 589 ! conserved hypothetical protein
STY0975 [imported] - Salmonella enterica su

1
E(R,Q,S,H,Y,E) (G,S,Q,I,T,N,P) (D,F,T,S,P,L,I) (N,S,K,N,T,D) (N,D,K,T,E) (L,F,R,I) (T,
S,N)R(I,V,A)

E(E) (G) (F) (D) (D) (F) (T)R(V)
439: IEQLD EEGFDDFTRV RELLG

T30186 ck: 3609 len: 543 ! hypothetical protein 8 - Shewanella
sp.

1
E(R,Q,S,H,Y,E) (G,S,Q,I,T,N,P) (D,F,T,S,P,L,I) (N,S,K,N,T,D) (N,D,K,T,E) (L,F,R,I) (T,
S,N)R(I,V,A)

E(Q) (Q) (L) (K) (D) (F) (S)R(A)
30: DVQVM EQQLKDFSRA CYVVN

B90396 ck: 6793 len: 220 ! hypothetical protein SSO2259
[imported] - Sulfolobus solfataricus

1
E(R,Q,S,H,Y,E) (G,S,Q,I,T,N,P) (D,F,T,S,P,L,I) (N,S,K,N,T,D) (N,D,K,T,E) (L,F,R,I) (T,
S,N)R(I,V,A)

E(E) (I) (I) (N) (D) (L) (S)R(I)
137: LINHW EEIINDLSRI DLTNF

A89606 ck: 9797 len: 692 ! protein F18G5.4 [imported] -
Caenorhabditis elegans

1
E(R,Q,S,H,Y,E) (G,S,Q,I,T,N,P) (D,F,T,S,P,L,I) (N,S,K,N,T,D) (N,D,K,T,E) (L,F,R,I) (T,
S,N)R(I,V,A)

E(E) (I) (L) (D) (E) (R) (S)R(I)

653: QHEIR EEILDERSRI QCSGN

Databases searched:

NBRF, Release 78.0, Released on 24Nov2003, Formatted on 25Nov2003

Total finds:	11
Total length:	96,191,526
Total sequences:	283,366
CPU time:	03:00.42

Blank

Swiss Prot
& SPTREMBL

! FINDPATTERNS on swp:* allowing 0 mismatches

! 1

query < E(R,Q,S,H,Y,E) (G,S,Q,I,T,N,P) (D,F,T,S,P,L,I) (N,S,K,N,T,D) (N,D,K,T,E) (L,F,R,I) (T,
S,N)R(I,V,A) October 5, 2004 12:52 ..

use ID
to match - FA10_HUMAN ck: 2459 len: 488 ! P00742 homo sapiens (human).
coagulation factor x precursor (ec 3.4.21.6)

sequence
to citation

1

E(R,Q,S,H,Y,E) (G,S,Q,I,T,N,P) (D,F,T,S,P,L,I) (N,S,K,N,T,D) (N,D,K,T,E) (L,F,R,I) (T,
S,N)R(I,V,A) - query

226: NQTQP
residue #

E(R) (G) (D) (N) (N) (L) (T) R(I) - matched portion of query
ERGDNNLTTRI VGGQE
matching region of db sequence

P100_HSV6U ck: 500 len: 870 ! Q00701 human herpesvirus (type 6 /
strain uganda-1102) (hhv6). large struc

1

E(R,Q,S,H,Y,E) (G,S,Q,I,T,N,P) (D,F,T,S,P,L,I) (N,S,K,N,T,D) (N,D,K,T,E) (L,F,R,I) (T,
S,N)R(I,V,A)

E(S) (Q) (F) (S) (D) (I) (N) R(A)
120: NKEKF ESQFSDINRA LLRLG

PFDA_METKA ck: 9337 len: 157 ! Q8tuy7 methanopyrus kandleri.
prefoldin alpha subunit (gimc alpha subunit)

1

E(R,Q,S,H,Y,E) (G,S,Q,I,T,N,P) (D,F,T,S,P,L,I) (N,S,K,N,T,D) (N,D,K,T,E) (L,F,R,I) (T,
S,N)R(I,V,A)

E(S) (S) (I) (S) (E) (L) (N) R(V)
39: QIDLI ESSISELNRV EETLK

VHS_VZVD ck: 3936 len: 455 ! P09275 varicella-zoster virus
(strain dumas) (vzv). virion host shutoff pr

1

E(R,Q,S,H,Y,E) (G,S,Q,I,T,N,P) (D,F,T,S,P,L,I) (N,S,K,N,T,D) (N,D,K,T,E) (L,F,R,I) (T,
S,N)R(I,V,A)

E(R) (G) (T) (S) (D) (L) (T) R(V)
107: AMNHD ERGTSDLTRV LLSSN

VPX_HV2D2 ck: 2803 len: 111 ! P15836 human immunodeficiency virus
type 2 (isolate d205,7) (hiv-2). vpx p

1

E(R,Q,S,H,Y,E) (G,S,Q,I,T,N,P) (D,F,T,S,P,L,I) (N,S,K,N,T,D) (N,D,K,T,E) (L,F,R,I) (T,
S,N)R(I,V,A)

E(R) (T) (I) (T) (E) (L) (N) R(V)
25: AFAWL ERTITELNRV AVNHL

YCAO_ECOLI ck: 8500 len: 586 ! P75838 escherichia coli.
hypothetical protein ycao. 10/2001

1
E(R,Q,S,H,Y,E) (G,S,Q,I,T,N,P) (D,F,T,S,P,L,I) (N,S,K,N,T,D) (N,D,K,T,E) (L,F,R,I) (T,
S,N)R(I,V,A)
E(E) (G) (F) (D) (D) (F) (T)R(V)
436: IEQLD EEGFDDFTRV RELLG

YZN4_CAEEL ck: 9797 len: 692 ! P54245 caenorhabditis elegans.
hypothetical acetylcholine receptor like pr

1
E(R,Q,S,H,Y,E) (G,S,Q,I,T,N,P) (D,F,T,S,P,L,I) (N,S,K,N,T,D) (N,D,K,T,E) (L,F,R,I) (T,
S,N)R(I,V,A)
E(E) (I) (L) (D) (E) (R) (S)R(I)
563: QHEIR EEILDERSRI QCSGN

O33907 ck: 3609 len: 543 ! O33907 shewanella sp. scrc-2738.
hypothetical protein. 10/2003

1
E(R,Q,S,H,Y,E) (G,S,Q,I,T,N,P) (D,F,T,S,P,L,I) (N,S,K,N,T,D) (N,D,K,T,E) (L,F,R,I) (T,
S,N)R(I,V,A)
E(Q) (Q) (L) (K) (D) (F) (S)R(A)
30: DVQVM EQQLKDFSRA CYVVN

Q9N4Q8 ck: 4608 len: 307 ! Q9n4q8 caenorhabditis elegans.
hypothetical protein. 10/2003

1
E(R,Q,S,H,Y,E) (G,S,Q,I,T,N,P) (D,F,T,S,P,L,I) (N,S,K,N,T,D) (N,D,K,T,E) (L,F,R,I) (T,
S,N)R(I,V,A)
E(S) (Q) (S) (N) (N) (L) (T)R(A)
213: LLKIH ESQSNNLTRA NRLAL

Q7XXN7 ck: 3447 len: 457 ! Q7xxn7 iris hollandica (dutch
iris). cytochrome p450. 10/2003

1
E(R,Q,S,H,Y,E) (G,S,Q,I,T,N,P) (D,F,T,S,P,L,I) (N,S,K,N,T,D) (N,D,K,T,E) (L,F,R,I) (T,
S,N)R(I,V,A)
E(S) (S) (D) (K) (E) (L) (N)R(V)
165: STLFR ESSDKELNRV ELKSK

Q9ESP3 ck: 4988 len: 1,851 ! Q9esp3 rattus norvegicus (rat).
mucin (fragment). 6/2003

1
E(R,Q,S,H,Y,E) (G,S,Q,I,T,N,P) (D,F,T,S,P,L,I) (N,S,K,N,T,D) (N,D,K,T,E) (L,F,R,I) (T,
S,N)R(I,V,A)
E(S) (N) (S) (T) (T) (L) (S)R(V)

130: LRRGL ESNSTTLSRV IMKLD

Q80Z21 ck: 6085 len: 1,726 ! Q80z21 mus musculus (mouse).
secreted gel-forming mucin (fragment). 10/200

1

E(R,Q,S,H,Y,E) (G,S,Q,I,T,N,P) (D,F,T,S,P,L,I) (N,S,K,N,T,D) (N,D,K,T,E) (L,F,R,I) (T,
S,N)R(I,V,A)

E(S) (N) (T) (T) (T) (L) (S)R(V)

126: LRRVQ ESNTTTLSRV TMKLD

Q8JVB6 ck: 3063 len: 895 ! Q8jvb6 helminthosporium victoriae
145s virus. putative capsid protein. 10/

1

E(R,Q,S,H,Y,E) (G,S,Q,I,T,N,P) (D,F,T,S,P,L,I) (N,S,K,N,T,D) (N,D,K,T,E) (L,F,R,I) (T,
S,N)R(I,V,A)

E(H) (I) (I) (N) (N) (R) (T)R(V)

552: PHDRV EHIINNRTRV TLSAR

Q76636 ck: 1892 len: 111 ! Q76636 human immunodeficiency virus
2. vpx protein. 6/2003

1

E(R,Q,S,H,Y,E) (G,S,Q,I,T,N,P) (D,F,T,S,P,L,I) (N,S,K,N,T,D) (N,D,K,T,E) (L,F,R,I) (T,
S,N)R(I,V,A)

E(R) (T) (I) (T) (E) (L) (N)R(V)

25: AFDWL ERTITELNRV AVNHL

Q7SNM2 ck: 973 len: 111 ! Q7snm2 human immunodeficiency virus
2. vpx protein. 10/2003

1

E(R,Q,S,H,Y,E) (G,S,Q,I,T,N,P) (D,F,T,S,P,L,I) (N,S,K,N,T,D) (N,D,K,T,E) (L,F,R,I) (T,
S,N)R(I,V,A)

E(R) (T) (I) (T) (E) (L) (N)R(V)

25: AFEWL ERTITELNRV AVNHL

Q8ZQC7 ck: 8250 len: 586 ! Q8zqc7 salmonella typhimurium.
putative cytoplasmic protein. 3/2003

1

E(R,Q,S,H,Y,E) (G,S,Q,I,T,N,P) (D,F,T,S,P,L,I) (N,S,K,N,T,D) (N,D,K,T,E) (L,F,R,I) (T,
S,N)R(I,V,A)

E(E) (G) (F) (D) (D) (F) (T)R(V)

436: IEQLD EEGFDDFTRV RELLG

Q8XEA9 ck: 2189 len: 589 ! Q8xea9 escherichia coli o157:h7.
orf, hypothetical protein. 6/2003

1
E(R,Q,S,H,Y,E) (G,S,Q,I,T,N,P) (D,F,T,S,P,L,I) (N,S,K,N,T,D) (N,D,K,T,E) (L,F,R,I) (T,
S,N)R(I,V,A)

E(E) (G) (F) (D) (D) (F) (T) R(V)
439: IEQLD EEGFDDFTRV RELLG

Q8FJB9 ck: 8041 len: 597 ! Q8fjb9 escherichia coli o6.
hypothetical protein ycao. 6/2003

1
E(R,Q,S,H,Y,E) (G,S,Q,I,T,N,P) (D,F,T,S,P,L,I) (N,S,K,N,T,D) (N,D,K,T,E) (L,F,R,I) (T,
S,N)R(I,V,A)

E(E) (G) (F) (D) (D) (F) (T) R(V)
447: IEQLD EEGFDDFTRV RELLG

Q897L2 ck: 3002 len: 375 ! Q897l2 clostridium tetani.
aminotransferase cobd. 10/2003

1
E(R,Q,S,H,Y,E) (G,S,Q,I,T,N,P) (D,F,T,S,P,L,I) (N,S,K,N,T,D) (N,D,K,T,E) (L,F,R,I) (T,
S,N)R(I,V,A)

E(S) (I) (L) (N) (N) (L) (N) R(V)
149: MDYDY ESILNNLN RV DGIIL

Q88HC5 ck: 1872 len: 528 ! Q88hc5 pseudomonas putida (strain
kt2440). conserved hypothetical protein.

1
E(R,Q,S,H,Y,E) (G,S,Q,I,T,N,P) (D,F,T,S,P,L,I) (N,S,K,N,T,D) (N,D,K,T,E) (L,F,R,I) (T,
S,N)R(I,V,A)

E(H) (I) (L) (D) (N) (L) (S) R(A)
57: ALEHV EHILDNLSRA ANSL

Q8Z806 ck: 3304 len: 589 ! Q8z806 salmonella typhi.
hypothetical protein sty0975. 6/2003

1
E(R,Q,S,H,Y,E) (G,S,Q,I,T,N,P) (D,F,T,S,P,L,I) (N,S,K,N,T,D) (N,D,K,T,E) (L,F,R,I) (T,
S,N)R(I,V,A)

E(E) (G) (F) (D) (D) (F) (T) R(V)
439: IEQLD EEGFDDFTRV RELLG

Q83LP5 ck: 8590 len: 597 ! Q83lp5 shigella flexneri. orf,
conserved hypothetical protein. 10/2003

1
E(R,Q,S,H,Y,E) (G,S,Q,I,T,N,P) (D,F,T,S,P,L,I) (N,S,K,N,T,D) (N,D,K,T,E) (L,F,R,I) (T,
S,N)R(I,V,A)

E(E) (G) (F) (D) (D) (F) (T) R(V)
447: IEQLD EEGFDDFTRV RELLG

Q7WM45 ck: 7664 len: 267 ! Q7wm45 bordetella bronchiseptica
(alcaligenes bronchisepticus). putative a

1

E(R,Q,S,H,Y,E) (G,S,Q,I,T,N,P) (D,F,T,S,P,L,I) (N,S,K,N,T,D) (N,D,K,T,E) (L,F,R,I) (T,
S,N)R(I,V,A)

E(R) (I) (I) (D) (D) (L) (N)R(A)

198: MIANI ERIIDDLNRA GQTFL

Q7W8I3 ck: 7664 len: 267 ! Q7w8i3 bordetella parapertussis.
putative amino acid abc transport system,

1

E(R,Q,S,H,Y,E) (G,S,Q,I,T,N,P) (D,F,T,S,P,L,I) (N,S,K,N,T,D) (N,D,K,T,E) (L,F,R,I) (T,
S,N)R(I,V,A)

E(R) (I) (I) (D) (D) (L) (N)R(A)

198: MIANI ERIIDDLNRA GQTFL

Q7UY73 ck: 3909 len: 1,080 ! Q7uy73 rhodopirellula baltica.
cation/multidrug efflux pump. 10/2003

1

E(R,Q,S,H,Y,E) (G,S,Q,I,T,N,P) (D,F,T,S,P,L,I) (N,S,K,N,T,D) (N,D,K,T,E) (L,F,R,I) (T,
S,N)R(I,V,A)

E(R) (G) (F) (S) (E) (I) (N)R(V)

842: AEINL ERGFSEINRV DQQRS

Q7UD35 ck: 8654 len: 586 ! Q7ud35 shigella flexneri.
hypothetical protein ycao. 10/2003

1

E(R,Q,S,H,Y,E) (G,S,Q,I,T,N,P) (D,F,T,S,P,L,I) (N,S,K,N,T,D) (N,D,K,T,E) (L,F,R,I) (T,
S,N)R(I,V,A)

E(E) (G) (F) (D) (D) (F) (T)R(V)

436: IEQLD EEGFDDFTRV RELLG

Q97WG2 ck: 6793 len: 220 ! Q97wg2 sulfolobus solfataricus.
hypothetical protein sso2259. 6/2003

1

E(R,Q,S,H,Y,E) (G,S,Q,I,T,N,P) (D,F,T,S,P,L,I) (N,S,K,N,T,D) (N,D,K,T,E) (L,F,R,I) (T,
S,N)R(I,V,A)

E(E) (I) (I) (N) (D) (L) (S)R(I)

137: LINHW EEIINDLSRI DLTNF

Databases searched:

SWISS-PROT, Release 42.7, Released on 15Dec2003, Formatted on 15Dec2003

SPTREMBL, Release 25.0, Released on 17Oct2003, Formatted on 18Oct2003

Total finds: 27
Total length: 367,588,357
Total sequences: 1,158,722

CPU time: 12:16.70